[MC-SMP]:
Session Multiplex Protocol Specification

Intellectual Property Rights Notice for Open Specifications Documentation

- **Technical Documentation.** Microsoft publishes Open Specifications documentation for protocols, file formats, languages, standards as well as overviews of the interaction among each of these technologies.
- **Copyrights.** This documentation is covered by Microsoft copyrights. Regardless of any other terms that are contained in the terms of use for the Microsoft website that hosts this documentation, you may make copies of it in order to develop implementations of the technologies described in the Open Specifications and may distribute portions of it in your implementations using these technologies or your documentation as necessary to properly document the implementation. You may also distribute in your implementation, with or without modification, any schema, IDL's, or code samples that are included in the documentation. This permission also applies to any documents that are referenced in the Open Specifications.
- **No Trade Secrets.** Microsoft does not claim any trade secret rights in this documentation.
- **Patents.** Microsoft has patents that may cover your implementations of the technologies described in the Open Specifications. Neither this notice nor Microsoft's delivery of the documentation grants any licenses under those or any other Microsoft patents. However, a given Open Specification may be covered by Microsoft's Open Specification Promise (available here: http://www.microsoft.com/interop/osp) or the Community Promise (available here: http://www.microsoft.com/interop/cp/default.mspx). If you would prefer a written license, or if the technologies described in the Open Specifications are not covered by the Open Specifications Promise or Community Promise, as applicable, patent licenses are available by contacting iplg@microsoft.com.
- **Trademarks.** The names of companies and products contained in this documentation may be covered by trademarks or similar intellectual property rights. This notice does not grant any licenses under those rights.
- **Fictitious Names.** The example companies, organizations, products, domain names, e-mail addresses, logos, people, places, and events depicted in this documentation are fictitious. No association with any real company, organization, product, domain name, email address, logo, person, place, or event is intended or should be inferred.

Reservation of Rights. All other rights are reserved, and this notice does not grant any rights other than specifically described above, whether by implication, estoppel, or otherwise.

Tools. The Open Specifications do not require the use of Microsoft programming tools or programming environments in order for you to develop an implementation. If you have access to Microsoft programming tools and environments you are free to take advantage of them. Certain Open Specifications are intended for use in conjunction with publicly available standard specifications and network programming art, and assumes that the reader either is familiar with the aforementioned material or has immediate access to it.

Revision Summary

Date	Revision History	Revision Class	Comments
08/10/2007	0.1	Major	Initial Availability
09/28/2007	0.2	Minor	Updated the technical content.
10/23/2007	0.2.1	Editorial	Revised and edited the technical content.
11/30/2007	0.2.2	Editorial	Revised and edited the technical content.
01/25/2008	0.2.3	Editorial	Revised and edited the technical content.
03/14/2008	0.2.4	Editorial	Revised and edited the technical content.
05/16/2008	0.2.5	Editorial	Revised and edited the technical content.
06/20/2008	0.3	Minor	Updated the technical content.
07/25/2008	0.3.1	Editorial	Revised and edited the technical content.
08/29/2008	1.0	Major	Updated and revised the technical content.
10/24/2008	1.0.1	Editorial	Revised and edited the technical content.
01/16/2009	1.0.2	Editorial	Revised and edited the technical content.
02/27/2009	1.0.3	Editorial	Revised and edited the technical content.
04/10/2009	1.0.4	Editorial	Revised and edited the technical content.
05/22/2009	2.0	Major	Updated and revised the technical content.
07/02/2009	2.0.1	Editorial	Revised and edited the technical content.
08/14/2009	2.0.2	Editorial	Revised and edited the technical content.
09/25/2009	2.1	Minor	Updated the technical content.
11/06/2009	3.0	Major	Updated and revised the technical content.
12/18/2009	3.0.1	Editorial	Revised and edited the technical content.
01/29/2010	3.1	Minor	Updated the technical content.
03/12/2010	4.0	Major	Updated and revised the technical content.
04/23/2010	5.0	Major	Updated and revised the technical content.
06/04/2010	5.0.1	Editorial	Revised and edited the technical content.
07/16/2010	5.0.1	No change	No changes to the meaning, language, or formatting of the technical content.
08/27/2010	5.0.1	No change	No changes to the meaning, language, or formatting of

Date	Revision History	Revision Class	Comments
			the technical content.
10/08/2010	6.0	Major	Significantly changed the technical content.
11/19/2010	7.0	Major	Significantly changed the technical content.
01/07/2011	8.0	Major	Significantly changed the technical content.
02/11/2011	8.0	No change	No changes to the meaning, language, or formatting of the technical content.

Contents

1 Introduction

This document specifies the Session Multiplex Protocol (SMP). SMP is an application-layer protocol that provides **session** management capabilities between a database **client** and a database **server**. Specifically, SMP enables multiple logical client connections to a single server over a lower-layer transport connection.

1.1 Glossary

The following terms are defined in [MS-GLOS]:

client
little-endian
server
session

The following terms are specific to this document:

Multiple Active Result Sets (MARS): A feature in SQL Server that allows applications to have more than one pending request per connection. Further information may be obtained from SQL Server Books Online ([MSDN-MARS]).

peer: The entity on either end of an established SMP session.

receiver: The entity that is receiving information from its **peer**. Both client and server can be receivers.

recycle: A process where SMP releases a **Session object**, such that the session identifier (SID) in use is made available again for a new session.

sender: The entity that is sending information to its **peer**. Both client and server can be senders.

session identifier (SID): A unique value provided by the **SID** field of a SYN packet to each session established over an SMP connection.

Session object: An instance of SMP created by a SYN packet, which corresponds to the SESSION ESTABLISHED state (section 3.1) and is designated by a unique session identifier (SID).

Session variable: Members of a **Session object** instance, which contain data to facilitate various SMP operations, such as messaging, event processing, and packet flow control.

Tabular Data Stream (TDS): An application-level protocol that is used by SQL Server to facilitate requests and responses between a database **server** and **client**, as specified in [MS-TDS].

Virtual Interface Architecture (VIA): A high-speed interconnect requiring special hardware and drivers provided by third parties, as specified in [VIA].

window: The number of SMP DATA packets that can be sent in the current **sender** state, per the SMP flow control algorithm that facilitates fairness among SMP sessions. This value is recalculated whenever a packet is sent or received.

MAY, SHOULD, MUST, SHOULD NOT, MUST NOT: These terms (in all caps) are used as described in [RFC2119]. All statements of optional behavior use either MAY, SHOULD, or SHOULD NOT.

1.2 References

1.2.1 Normative References

We conduct frequent surveys of the normative references to assure their continued availability. If you have any issue with finding a normative reference, please contact dochelp@microsoft.com. We will assist you in finding the relevant information. Please check the archive site, http://msdn2.microsoft.com/en-us/library/E4BD6494-06AD-4aed-9823-445E921C9624, as an additional source.

[MS-DTYP] Microsoft Corporation, "Windows Data Types", January 2007.

[RFC793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981, http://www.ietf.org/rfc/rfc0793.txt

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997, http://www.ietf.org/rfc/rfc2119.txt

[RFC2246] Dierks, T., and Allen, C., "The TLS Protocol Version 1.0", RFC 2246, January 1999, http://www.ietf.org/rfc/rfc2246.txt

[SSL3] Netscape, "SSL 3.0 Specification", http://tools.ietf.org/html/draft-ietf-tls-ssl-version3-00

If you have any trouble finding [SSL3], please check here.

1.2.2 Informative References

[MS-GLOS] Microsoft Corporation, "Windows Protocols Master Glossary", March 2007.

[MSDN-MARS] Microsoft Corporation, "Multiple Active Result Sets (MARS) in SQL Server 2005", http://msdn.microsoft.com/en-us/library/ms345109.aspx

[MS-TDS] Microsoft Corporation, "Tabular Data Stream Protocol Specification", February 2008.

[PIPE] Microsoft Corporation, "Named Pipes", http://msdn.microsoft.com/en-us/library/aa365590.aspx

[VIA] Intel Corporation, "Intel Virtual Interface (VI) Architecture Developer's Guide", September 1998, ftp://download.intel.com/design/servers/vi/

1.3 Overview

SMP is an application protocol that facilitates session management by providing a mechanism to create multiple lightweight communication channels (sessions) over a lower-layer transport connection. SMP does this by multiplexing data streams from different sessions on top of a single reliable stream-oriented transport.

SMP is beneficial in situations where database connections from the client and server are synchronous. In this context, "synchronous" means that the client application can only have one outstanding command or transaction per connection. Rather than incur the expense of creating multiple connections to the server, SMP is capable of simultaneously executing multiple database queries over a single connection.

SMP provides the following:

- The ability to interleave data from several different sessions and preserve message boundaries.
- A sliding window-based flow-control mechanism to facilitate fairness among sessions.

Note SMP is defined as a transport-independent mechanism. It relies on an underlying transport mechanism such as Transmission Control Protocol (TCP) (specified in [RFC793]) to ensure byte alignment, loss detection and recovery, and reliable in-order delivery. The scheduling algorithm that enforces fairness between sessions is an implementation issue for the application that implements SMP.

The following diagram shows typical SMP communication flow for an arbitrary session.

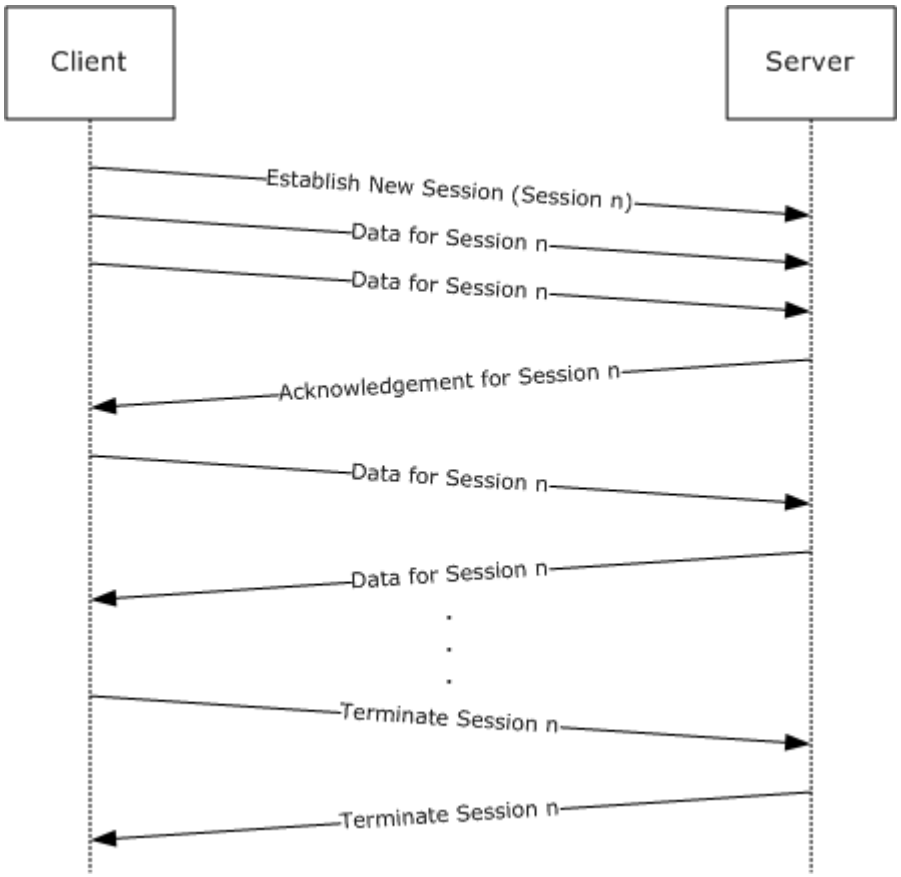


Figure 1: Example of a communication flow in SMP

1.4 Relationship to Other Protocols

SMP depends on an underlying reliable stream-oriented network transport. Optionally, Transport Layer Security (TLS)/Secure Sockets Layer (SSL) ([RFC2246] and [SSL3]) can be inserted between SMP and the transport layer to provide data protection.

The **Tabular Data Stream (TDS)** protocol, as specified in [MS-TDS], depends on SMP when the **Multiple Active Result Sets (MARS)** feature is specified ([MSDN-MARS]). TDS is an example of a higher-layer protocol for SMP. This dependency is illustrated in the following diagram.

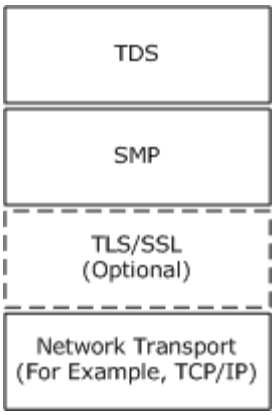


Figure 2: Protocol relationship

1.5 Prerequisites/Preconditions

It is assumed throughout this document that the client has already discovered the server and established a network transport connection.

1.6 Applicability Statement

SMP is used appropriately to facilitate the multiplexing of several sessions over a single reliable lower-layer transport connection where network or local connectivity is available.

1.7 Versioning and Capability Negotiation

No other version of SMP exists than the protocol that is described in this specification. Additional details follow.

Supported transports: SMP can be implemented on top of any reliable transport mechanism, as specified in section 2.1.

Protocol versions: SMP supports the SMP 1.0 version, which is the only version of SMP available, as defined in section 2.2.

Security and authentication methods: SMP does not provide or support any security or authentication methods.

Localization: SMP does not provide any localization-specific features.

Capability negotiation: SMP does not support capability negotiation.

1.8 Vendor-Extensible Fields

There are no vendor-extensible fields.

1.9 Standards Assignments

There are no standards assignments for SMP.

2 Messages

All integer fields are represented in **little-endian** byte order. This protocol references commonly used data types as defined in [MS-DTYP].

2.1 Transport

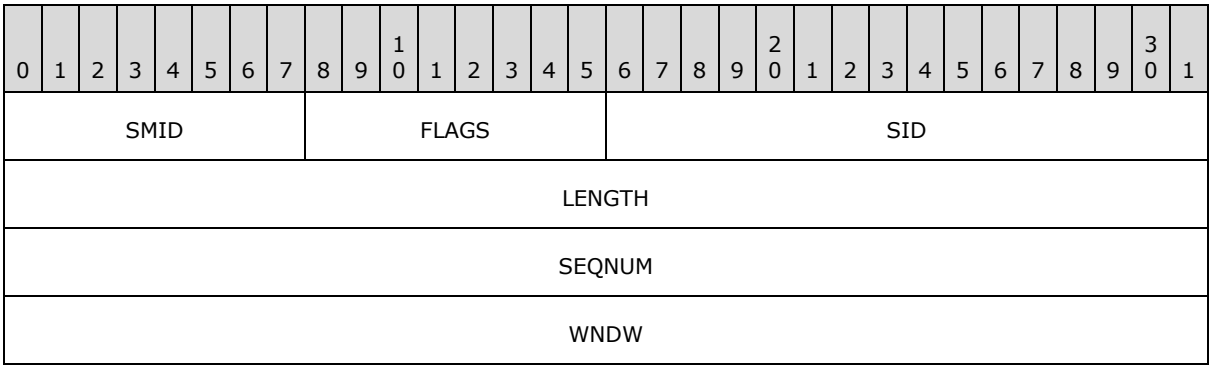
SMP is a simple protocol that is layered above existing reliable transport mechanisms, such as TCP ([RFC793]), named pipes ([PIPE]), or **Virtual Interface Architecture** ([VIA]). SMP enables the creation of multiple sessions over a single connection. SMP is defined as a transport-independent mechanism.

2.2 Message Syntax

All SMP packets consist of a 16-byte header followed by an optional data payload, depending on the packet type.

2.2.1 Header

The 16-byte SMP header has the following format.



SMID (1 byte): This unsigned integer is the SMP packet identifier and MUST always be assigned the value 0x53. This field indicates that the packet is an SMP packet, which helps to distinguish it from other protocol packets.

FLAGS (1 byte): This unsigned integer value contains the control flags, as defined in section 2.2.1.1.

SID (2 bytes): This unsigned integer is the session identifier. This value is a unique identifier for each session that is multiplexed over this connection.

LENGTH (4 bytes): This unsigned integer specifies the length, in bytes (including the header), of the SMP packet.

SEQNUM (4 bytes): This unsigned integer is the SMP sequence number for this packet in the session. The first DATA packet in each session MUST have a SEQNUM value of 0x00000001. For every DATA packet thereafter, this integer MUST monotonically increase by a value of 1 up to 0xffffffff, and then wraps back to a starting value of 0x00000000. Sequence numbers MUST only be incremented for DATA packets. For the ACK packet type, the sequence number MUST remain stable. For the FIN packet type, the sequence number SHOULD remain stable. For the SYN packet type, the sequence number SHOULD be 0x00000000.

WNDW (4 bytes): This unsigned integer indicates the maximum SEQNUM value permitted for a receive packet.

Note The difference between the values of the WNDW field of a received packet and the SEQNUM field of the last sent packet is the available send window size. Any subsequent packets that are sent MUST NOT contain a SEQNUM value that is greater than the value of the WNDW field of the last received packet.

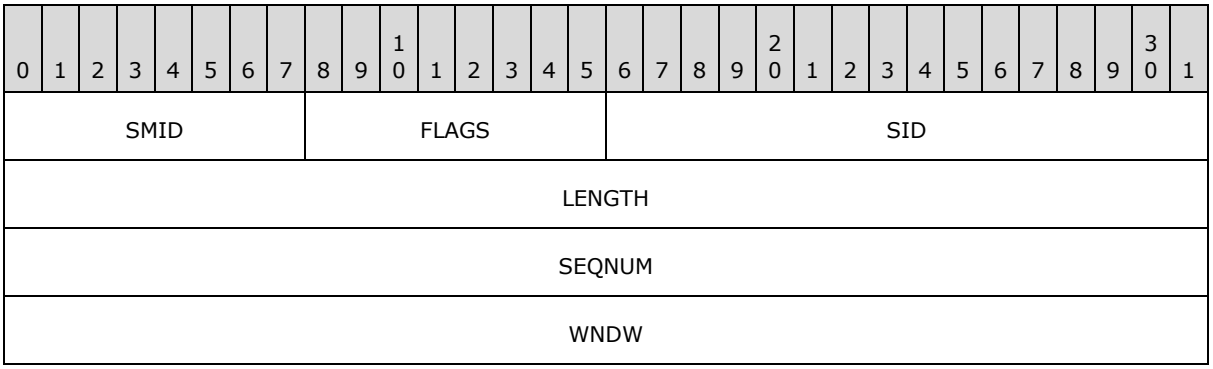
2.2.1.1 Control Flags

The control flag is 1 byte after the **SMID** field and indicates the type of the packet. Only DATA packets have payload data. The **sender** MUST NOT send a combination of flags in the same packet. For example, a **FLAGS** field value of 0x06 (ACK plus FIN) is an invalid value.

Value	Meaning
SYN 0x01	Indicates that a new connection is to be established (see SYN packet). The session ID for the session is the number that is stored in the **SID** field.
ACK 0x02	Informs the **peer** about a change in window size when consecutive unanswered DATA packets are received (see ACK packet).
FIN 0x04	Indicates that the sending entity will no longer use the session to send data.
DATA 0x08	Indicates that the packet carries user data after the header (see DATA packet).

2.2.2 SYN Packet

The SYN packet is sent to indicate that a new connection is to be established. The ID for the session is the number that is stored in the **SID** field of the SYN packet.



SMID (1 byte): See section 2.2.1 for a description of the **SMID** field.

FLAGS (1 byte): This unsigned integer contains control flags that identify this packet as a SYN packet. The value of the **FLAGS** field MUST be 0x01. See section 2.2.1.1 for details.

SID (2 bytes): All subsequent packets in this session MUST use this identifier. See section 2.2.1 for a description of the **SID** field.

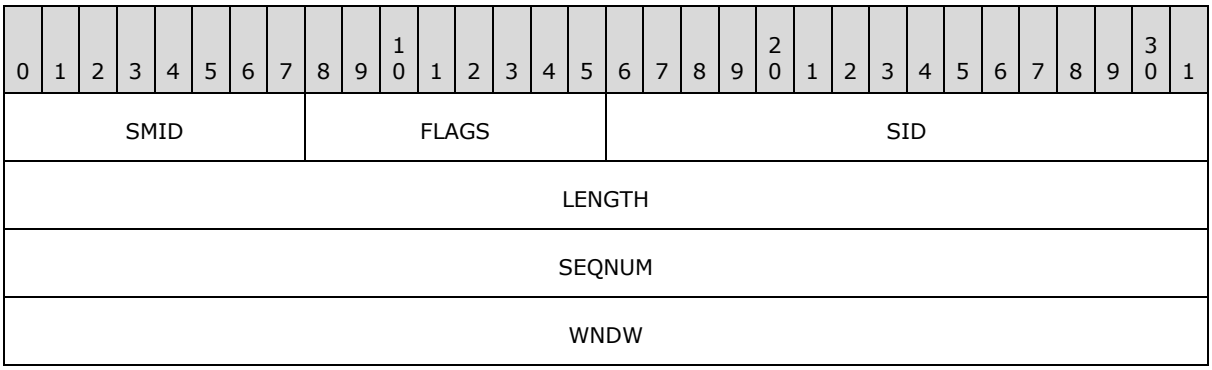
LENGTH (4 bytes): The value of this field MUST be 0x00000010. See section 2.2.1 for a description of the **LENGTH** field.

SEQNUM (4 bytes): The value of this field SHOULD be 0x00000000. See section 2.2.1 for a description of the **SEQNUM** field.

WNDW (4 bytes): See section 2.2.1 for a description of the **WNDW** field.

2.2.3 ACK Packet

The ACK packet updates the peer by changing the peer's send window size when several consecutive unanswered DATA packets are received. For example, with a send window size of 4 (the value of the **WNDW** field of the sender's last received packet is equal to 0x00000004, and the value of the **SEQNUM** field of the sender's next sent packet will be equal to 0x00000001), if the sender has 5 packets to pass to the **receiver** for a single request, then after 4 packets the sender will wait until it receives an ACK packet with an updated value for the **WNDW** field before it can transmit additional packets. After the receiver has processed at least one of the packets, the receiver can send the sender an ACK packet containing an updated **WNDW** field value, which allows the sender to send the final packet and complete the request.



SMID (1 byte): See section 2.2.1 for a description of the **SMID** field.

FLAGS (1 byte): This unsigned integer contains control flags that identify this packet as an ACK packet. The value of the **FLAGS** field value MUST be 0x02.

SID (2 bytes): See section 2.2.1 for a description of the **SID** field. This MUST be the value that was set in the SYN packet (when the session was opened).

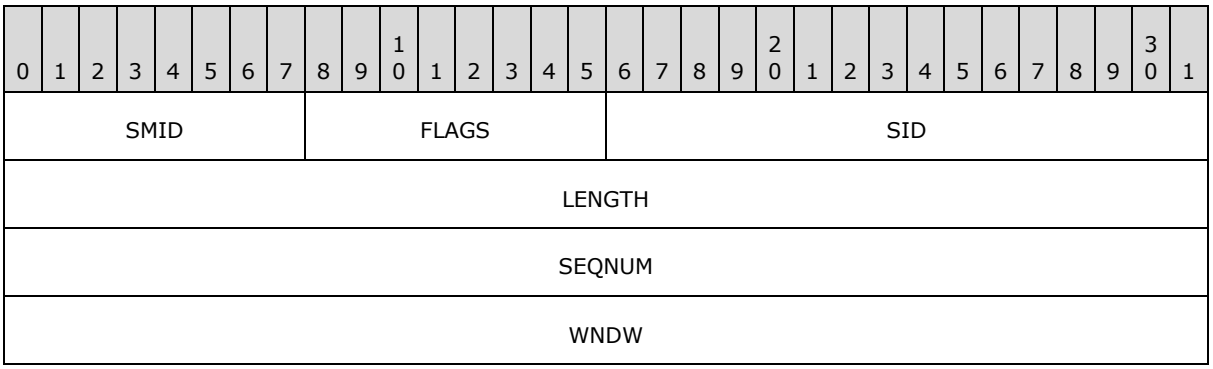
LENGTH (4 bytes): See section 2.2.1 for a description of the **LENGTH** field. The value of this field MUST be 0x00000010.

SEQNUM (4 bytes): See section 2.2.1 for a description of the **SEQNUM** field.

WNDW (4 bytes): See section 2.2.1 for a description of the **WNDW** field.

2.2.4 FIN Packet

The FIN packet is sent to indicate that the sending entity will no longer use the session to send or receive data.



SMID (1 byte): See section 2.2.1 for a description of the **SMID** field.

FLAGS (1 byte): This unsigned integer contains control flags that identify this packet as a FIN packet. The value of the **FLAGS** field MUST be 0x04.

SID (2 bytes): The **SID** field MUST be set to the value that was set when the session was opened. See section 2.2.1 for a description of the **SID** field.

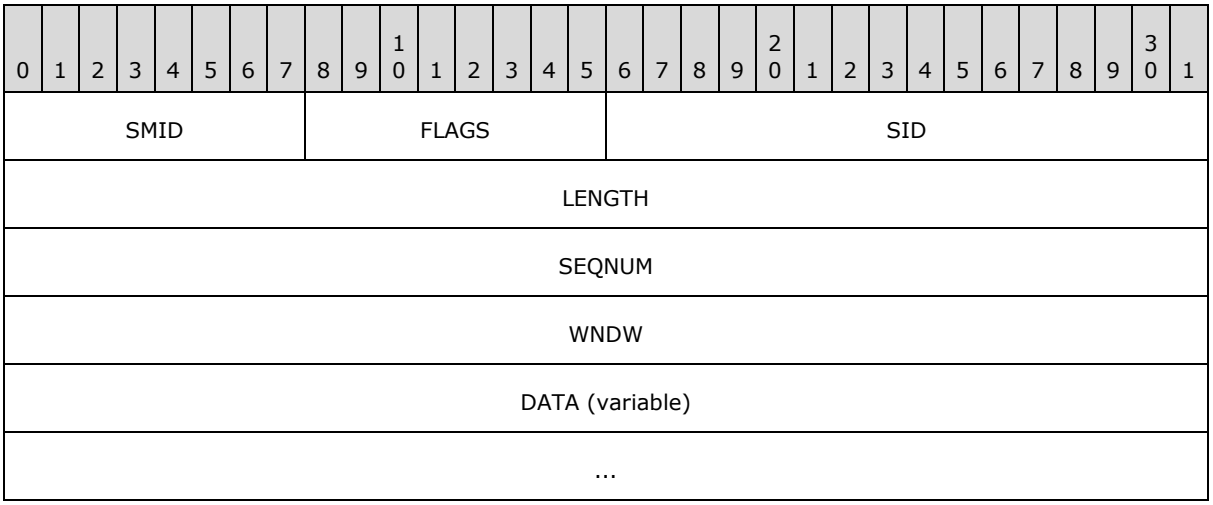
LENGTH (4 bytes): The value of the **LENGTH** field MUST be 0x00000010. See section 2.2.1 for a description of the **LENGTH** field.

SEQNUM (4 bytes): See section 2.2.1 for a description of the **SEQNUM** field.

WNDW (4 bytes): See section 2.2.1 for a description of the **WNDW** field.

2.2.5 DATA Packet

The DATA packet carries data in the **DATA** field, which follows the Header. The length of the **DATA** field is the total SMP packet length minus the SMP packet header length.



SMID (1 byte): See section 2.2.1 for a description of the **SMID** field.

FLAGS (1 byte): This unsigned integer contains control flags that identify this packet as a DATA packet. The value of the **FLAGS** field MUST be 0x08.

SID (2 bytes): The **SID** field MUST be set to the value that was set when the session was opened. See section 2.2.1 for a description of the **SID** field.

LENGTH (4 bytes): The value of the **LENGTH** field MUST be at least 0x00000010. See section 2.2.1 for a description of the **LENGTH** field.

SEQNUM (4 bytes): See section 2.2.1 for a description of the **SEQNUM** field.

WNDW (4 bytes): See section 2.2.1 for a description of the **WNDW** field.

DATA (variable): The **DATA** field contains the user data of the DATA packet. The size of the **DATA** field can be determined by subtracting the length of the Header (16 bytes) from the value of the **LENGTH** field. For example, a **LENGTH** value of 0x00000025 means the user data will be 21 bytes long.

3 Protocol Details

This section describes the important elements of the client and server software necessary to support SMP.

SMP is largely a symmetric protocol that obeys the same rules and semantics on both the client and the server. Therefore, descriptions of the client and server roles are both contained in section 3.1, where section 3.3.2.2 applies only to the client and section 3.2.4.1 applies only to the server.

3.1 Common Details

SMP MUST be layered on top of a reliable, in-order, connection-oriented transport layer such as TCP ([RFC793]), named pipes ([PIPE]), or Virtual Interface Architecture, as specified in [VIA].

After the transport connection is established, SMP initiation MUST be negotiated through other protocols, such as [MS-TDS]. SMP MUST be successfully initiated on both end points before SMP operations can begin. The shutdown sequence can be triggered either by the higher layer or by fatal events internal to SMP. The peer is notified of the shutdown when the transport connection is closed.

SMP incorporates the concept of a client and a server interacting during session establishment. A session MUST be initiated by the client (section 3.3.2.2). After SMP enters the SESSION ESTABLISHED state, both endpoints of the session can be used by the higher layer to send and receive data symmetrically, and therefore each can act as a sender and as a receiver. Either the client or the server can initiate connection termination by sending a FIN packet.

3.1.1 Abstract Data Model

This section describes a conceptual model of possible data organization that an implementation maintains to participate in this protocol. The described organization is provided to facilitate the explanation of how the protocol behaves. This document does not mandate that implementations adhere to this model as long as their external behavior is consistent with what is described in this document.

3.1.1.1 Session-Specific Structures

The following structures are required per SMP session. These structures are needed to implement the flow control algorithm and for connection management:

Note The dotted notation of the following list items indicates the structures of a **Session object** instance. For example, Session.SeqNumForSend refers to the SeqNumForSend variable of the Session object.

- **Session.SeqNumForSend**: A 32-bit unsigned integer that monotonically increases for every session DATA packet that is sent.
- **Session.HighWaterForSend**: A 32-bit unsigned integer that tracks the peer window that is obtained through the **WNDW** field in the received packet header.
- **Session.SeqNumForRecv**: A 32-bit unsigned integer that tracks the peer session sequence number obtained from the **SEQNUM** field in the DATA packet header. This number is used for comparison with the received packet.

- **Session.HighWaterForRecv**: A 32-bit unsigned integer that tracks the receiver's high-water mark of the receiver buffer window. It is used to set the value of the **WNDW** field of each sent packet.
- **Session.LastHighWaterForRecv**: A 32-bit unsigned integer that tracks the value of the **WNDW** field of the last sent packet. It is used to implement a selective ACK algorithm and is optional.
- **Session.ReceivePacketQueue**: A queue that buffers received packets.

3.1.1.2 Session States

The state of an SMP session has to be maintained. An SMP session can be in any one of several states, which are described here and in the "Server Details" and "Client Details" sections.

SESSION ESTABLISHED: The session is successfully established and both session endpoints can send and receive data. The SESSION ESTABLISHED state is reached as specified in section 3.3.2.2.

FIN RECEIVED: The client or server has received a FIN packet from its peer, indicating a request to close this session.

FIN SENT: The client or server has sent a FIN packet to its peer after the session is closed by the higher layer. The sending entity will also receive a FIN packet from its peer before entering the CLOSED state (section 3.1.4.4). However, if the transport connection will also be closed by the sending entity, it is unnecessary to wait to receive the FIN packet acknowledgement.

CLOSED: The session has been closed by the higher layer, either by closing the SMP session (section 3.1.4.4) or by shutting down the SMP connection (section 3.1.4.5), at which point data can no longer be sent or received.

3.1.2 Timers

In SMP, there are no timers. SMP assumes a reliable transport and the eventual delivery of messages. In the event of an error from the transport connection, SMP **recycles** all Session objects associated with the failed transport connection. Idle sessions are kept open until the higher layer closes them or an error in the transport connection occurs.

3.1.3 Initialization

3.1.3.1 Session-Specific Structure

Session-specific structures MUST be initialized with the values described in the table that follows.

Note The dotted notation in the table indicates the structures of a Session object instance. For example, Session.SeqNumForSend refers to the SeqNumForSend variable of the Session object.

Variable	Value
Session.SeqNumForSend	0
Session.HighWaterForSend	4
Session.SeqNumForRecv	0

Variable	Value
Session.HighWaterForRecv	4
Session.ReceivePacketQueue	Empty

If the delayed acknowledgment algorithm is used, as specified in section 3.1.5.2.3, Session.LastHighWaterforRecv will have a value of 4. Otherwise, the Session.LastHighWaterForRecv variable is not used.

3.1.4 Higher-Layer Triggered Events

3.1.4.1 Initialize by Higher Layer

The higher layer on both the client and server MUST initialize SMP on each end of the lower-layer transport connection before SMP can operate. After initialization, the client enters a CLOSED state and the server enters a LISTENING state.

3.1.4.2 Read by Higher Layer

The Read by Higher Layer event is triggered when the higher layer chooses to perform a read operation on arriving DATA packets. The following should occur when a DATA packet arrives:

- If the ReceivePacketQueue variable of the Session object is empty, the SMP layer MUST notify the higher layer once a DATA packet arrives, as described in section 3.1.5.1.1.
- If the ReceivePacketQueue variable of the Session object is not empty, then the SMP layer MUST retrieve only one packet from the ReceivePacketQueue and pass it to the higher layer. After the SMP layer passes the data to the higher layer, the HighWaterForRecv variable of the Session object is incremented by 1, and the SMP layer can send an ACK packet to the peer, as specified in section 3.1.5.2.3.

3.1.4.3 Higher Layer Initiates Sending of Data

This event is triggered when the higher layer initiates the sending of data over an SMP session.

Any packet that is sent MUST NOT contain a **SEQNUM** value higher than the value of the HighWaterForSend variable.

If a DATA packet cannot be sent to its peer because the value of the SeqNumForSend variable of the Session object is equal to the value of the HighWaterForSend variable of the Session object, the SMP layer MUST choose to perform one of the following two actions:

- Buffer the DATA packet in a local buffer and send it at a later time according to the flow control algorithm described in section 3.1.5.2.
- Block the higher layer until the DATA packet is sent according to the flow control algorithm described in section 3.1.5.2.

If the value of SeqNumForSend is less than that of HighWaterForSend, the SMP layer of the sender MUST send the DATA packet according to the flow control algorithm described in section 3.1.5.2.

3.1.4.4 Close by Higher Layer

The Close by Higher Layer event is triggered when the upper layer closes a session. When this happens, the following MUST occur:

1. If SMP is in the SESSION ESTABLISHED state, send the FIN packet and enter the FIN SENT state.
2. If SMP is in the FIN RECEIVED state, send the FIN packet to the peer, recycle the Session object, and then enter the CLOSED state.

Note The Session object cannot be recycled and the CLOSED state should not be entered until the SMP layer receives a FIN packet from its peer, as described in section 3.1.5.1.3. It is also important to both receive and send a FIN packet (the order does not matter) before entering the CLOSED state to prevent a new session from attempting to use an existing **session identifier** (SID).

3.1.4.5 Shutdown by Higher Layer

The Shutdown by Higher Layer event is triggered when the upper layer shuts down the SMP connection. When this occurs, all sessions MUST move from the CLOSED state to the END state and all associated data structures MUST be released.

3.1.5 Message Processing Events and Sequencing Rules

3.1.5.1 Receiving a Packet

The client or server MUST do the following when receiving a packet:

1. Parse the header of the received packet to get the value of the **SID** field.
2. If the Session object corresponding to the value of the **SID** field of the received packet does not exist and the value of the **FLAGS** field of the packet is not equal to 0x01 (a SYN packet), an error MUST be raised to the higher layer and the underlying transport connection MUST be closed.
3. If the Session object is located, an error MUST be raised to the higher layer and the underlying transport connection MUST be closed if any of the following conditions are not met:
   - The value of the **FLAGS** field in the received packet is equal to 0x02 (ACK packet), 0x04 (FIN packet), or 0x08 (DATA packet).
   - The value of the **WNDW** field of the received packet is greater than or equal to the value of the HighWaterForSend variable of the Session object.
   - The **SID** field of the received packet matches the SID of the Session object.
   - The value of the **SEQNUM** field is less than or equal to the value of the HighWaterForRecv variable of the Session object.
4. If the value of the **FLAGS** field is equal to 0x08 (a DATA packet), parse the packet to get the user data (**DATA** field) while using the value of the **LENGTH** field of the packet to facilitate the parse.

   Note The length of the **DATA** field will be equal to the overall packet **LENGTH** minus the length of the Header (16 bytes).

The sections that follow describe the processing of received DATA, ACK, and FIN packets. Processing of received SYN packets is covered in the server- and client-specific sections.

3.1.5.1.1 Receiving a DATA Packet

When a DATA packet is received in the SESSION ESTABLISHED state:

1. If a higher layer posted a receive, finish that receive with the data in the packet; otherwise, buffer the packet in the ReceivePacketQueue variable of the Session object.

2. If the value of the **WNDW** field of the DATA packet is greater than the value of the HighWaterForSend variable of the Session object, the receiver of the DATA packet MUST do the following:

   - If there are any packets waiting to be sent (section 3.1.4.3), then the SMP layer MUST send the packets to its peer, up to and including the value of the packet number defined by the **WNDW** field.

   - Set the value of the HighWaterForSend variable of the Session object equal to the value of the **WNDW** field of the DATA packet.

3. If the value of the **SEQNUM** field of the DATA packet is not equal to the value of the SeqNumForRecv variable of the Session object plus 1, an error MUST be raised to the higher layer and the underlying transport layer MUST be closed.

Note When a DATA packet is received in the FIN SENT state, the packet MUST be ignored.

Note When a DATA packet is received in the FIN RECEIVED state, an error SHOULD be raised to the higher layer and the underlying transport connection SHOULD be closed.

3.1.5.1.2 Receiving an ACK Packet

When an ACK packet is received, the following applies:

- If the value of the **WNDW** field of the ACK packet is greater than the value of the HighWaterForSend variable of the Session object, then the receiver of the ACK packet MUST do the following:

  - If there are any packets waiting to be sent, as specified in section 3.1.4.3, then the SMP layer MUST send the packets to its peer, up to and including the value defined by the **WNDW** field.

  - Set the value of the HighWaterForSend variable to that of the **WNDW** field.

- If an ACK packet is received in the FIN RECEIVED state, an error SHOULD be raised to the higher layer and the underlying transport connection SHOULD be closed.

- If the value of the **SEQNUM** field of the ACK packet is not equal to the value of the SeqNumForRecv variable of the Session object, an error MUST be raised to the higher layer and the underlying transport connection MUST be closed.

3.1.5.1.3 Receiving a FIN Packet

When a FIN packet is received, the following applies:

1. If SMP is in the SESSION ESTABLISHED state, then move into the FIN RECEIVED state.

2. If SMP is in the FIN SENT state, then recycle the Session object and move into the CLOSED state.

When a FIN packet is received in the FIN RECEIVED state, an error SHOULD be raised to the higher layer and the underlying transport connection SHOULD be closed.

If the value of the **SEQNUM** field of the FIN packet is not equal to the value of the SeqNumForRecv variable of the Session object, an error MAY be raised to the higher layer and the underlying transport connection MAY be closed.

3.1.5.2 Flow Control Algorithm

SMP provides a means for the receiver to govern the amount of data sent by the sender. This is achieved by returning a window with every ACK or DATA packet. The returned window indicates a range of acceptable sequence numbers beyond the last DATA packet that is successfully received. The window indicates an allowed number of DATA packets that the sender may transmit before receiving further permission.

Flow control involves the use of the following sender variables:

- Session.SeqNumForSend
- Session.HighWaterForSend

Flow control also involves the use of the following receiver variables:

- Session.SeqNumForRecv
- Session.HighWaterForRecv
- LastHighWaterForRecv

The sections that follow show the relationships of these variables in the sequence number space. The sequence number is a 32-bit unsigned integer that is allowed to wrap.

3.1.5.2.1 Session Variable Relationships for the Sender

1. The DATA packet MUST NOT be sent if the value of the SeqNumForSend variable of the Session object is equivalent to the value of the HighWaterForSend variable of the Session object.
2. Otherwise, the value of the **SEQNUM** field of the DATA packet MUST be set to the value of the SeqNumForSend variable plus 1, the DATA packet MUST be sent, and the value of the SeqNumForSend variable MUST then be incremented by 1.
3. Upon receiving a packet, the value of the HighWaterForSend variable MUST be set to the value of the **WNDW** field of the received packet.

Note The value of the send window size equals the value of the HighWaterForSend variable minus the value of the SeqNumForSend variable. The send window is considered closed when the value of the send window size is 0. The maximum send window size for the implementation described in this document is 4.

3.1.5.2.2 Session Variable Relationships for the Receiver

1. When the higher layer retrieves a DATA packet from a session endpoint, increment the HighWaterForRecv variable of the Session object by 1.
2. When sending a DATA packet, the value of the **WNDW** field of the packet MUST be set to the value of the HighWaterForRecv variable.
3. When receiving a DATA packet with a **SEQNUM** field value equivalent to the value of the SeqNumForRecv variable of the Session object plus 1 (and that value is less than or equal to the value of the HighWaterForRecv variable of the Session object), then the value of the

SeqNumForRecv variable MUST be set to the value of the **SEQNUM** field of the received DATA packet.

4. When receiving a DATA packet with a **SEQNUM** field that does not satisfy the condition specified above, an error MUST be raised to the higher layer.

5. When receiving a packet other than a DATA packet, the SeqNumForRecv variable MUST NOT be changed.

Note The algorithm described above ensures that, at any time, the value of the SeqNumForRecv variable is less than or equal to the value of the HighWaterForRecv variable. The receive window size equals the value of HighWaterForRecv minus the value of SeqNumForRecv.

3.1.5.2.3 Update Sender's HighWaterForSend Variable Using an ACK Packet

The ADM variable **HighWaterForSend** of the Session object is updated by receiving either a DATA packet or an ACK packet from the peer. The SMP layer MUST send ACK packets to facilitate flow control. There are several possible algorithms that can be used for sending ACK packets; this is an implementation choice. One example is to send an ACK packet for each DATA packet retrieved by the higher layer.<1>

3.1.6 Timer Events

There is no timer in SMP.

3.1.7 Other Local Events

In case of the following events, SMP MUST close the lower layer transport connection and an error MUST be raised to the higher layer:

1. The lower-layer transport disconnects.

2. A packet is received by a peer and does not follow the specifications outlined in section 2.

3.2 Server Details

The following state diagram illustrates the progress of a session during the lifetime of the server. The diagram is only a summary and does not represent the total specification; for example, it does not include error events and state changes within an established state.

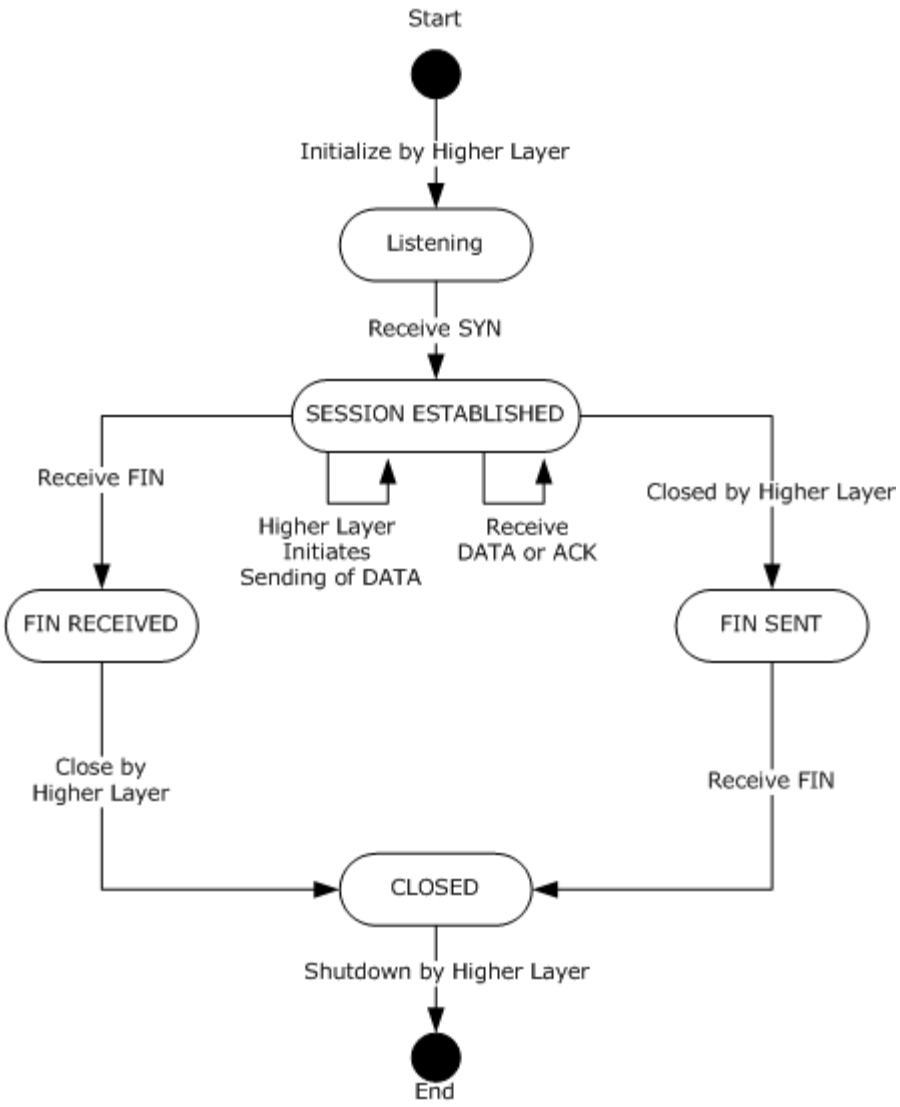


Figure 3: Session Multiplex Protocol server state machine

3.2.1 Initialization

On the server side, initialization of the Abstract Data Model described in the Common Details is performed when the upper layer makes a request to begin listening.

3.2.2 Higher-Layer Triggered Events

3.2.2.1 Initialize by Higher Layer

The higher layer on the server MUST initialize SMP on each end of the lower-layer transport connection before SMP can operate. After initialization, the server enters a LISTENING state.

3.2.3 Session States

In addition to the states specified in the Common Details, a Server Session may also be in the following state:

Listening: The server is ready for client connections.

3.2.4 Processing Events and Sequencing Rules

3.2.4.1 Receiving a SYN Packet

The following logic applies to the server only when receiving a SYN packet.

- Create a Session object using the value of the **SID** field of the received SYN packet and enter the SESSION ESTABLISHED state.
- If the value of the **SEQNUM** field of the SYN packet is not equal to the value of the SeqNumForRecv variable of the Session object, an error MAY be raised to the higher layer and the underlying transport connection MAY be closed.

Note If a SYN packet is received in the FIN RECEIVED state, an error SHOULD be raised to the higher layer and the underlying transport connection SHOULD be closed.

3.3 Client Details

The following state diagram illustrates the progress of a session during the lifetime of the client. The diagram is only a summary and does not represent the total specification; for example, it does not include error events and state changes within an established state.

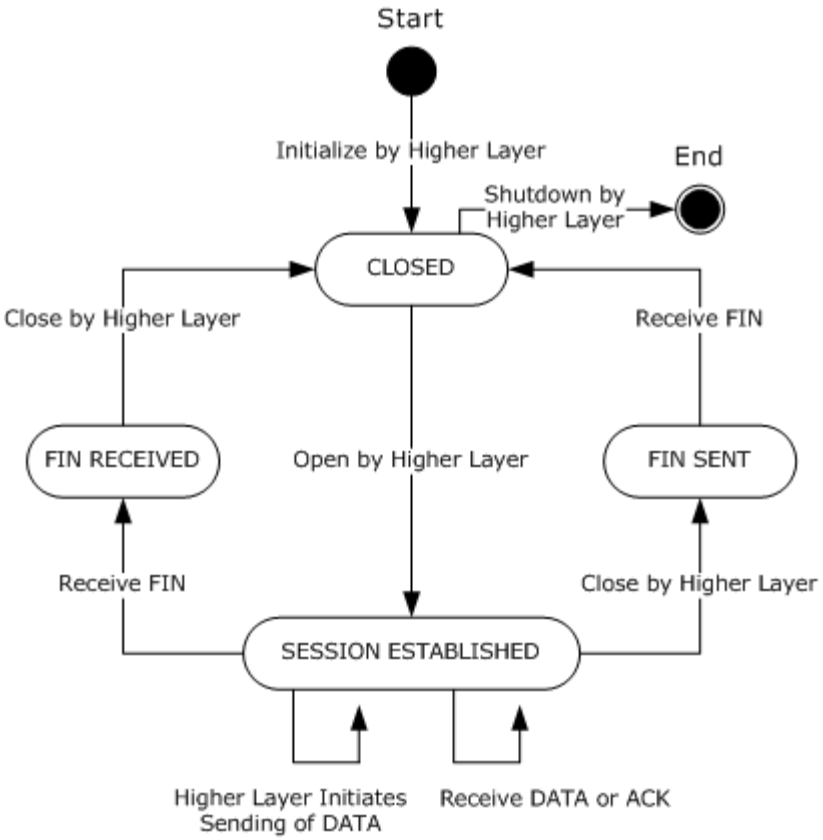


Figure 4: Session Multiplex Protocol client state machine

3.3.1 Initialization

On the client side, initialization of the Abstract Data Model described in the Common Details is performed when the upper layer makes a request for a new SMP session.

3.3.2 Higher-Layer Triggered Events

3.3.2.1 Initialize by Higher Layer

The higher layer on the client MUST initialize SMP on each end of the lower-layer transport connection before SMP can operate. After initialization, the client enters a CLOSED state.

3.3.2.2 Open by Higher Layer

The Open by Higher Layer event is triggered from the client side only. When the higher layer triggers this event, the SMP layer MUST:

1. Choose a unique session identifier (SID), as specified in section 2.2.1, for each session multiplexed over a lower-layer transport connection.

2. Send a SYN packet to the server.

Note The SYN packet creates a Session object, which is an instance of the SMP protocol containing **Session variables** that control protocol operation.

3. Enter into the SESSION ESTABLISHED state.

3.3.3 Processing Events and Sequencing Rules

3.3.3.1 Receiving a SYN Packet

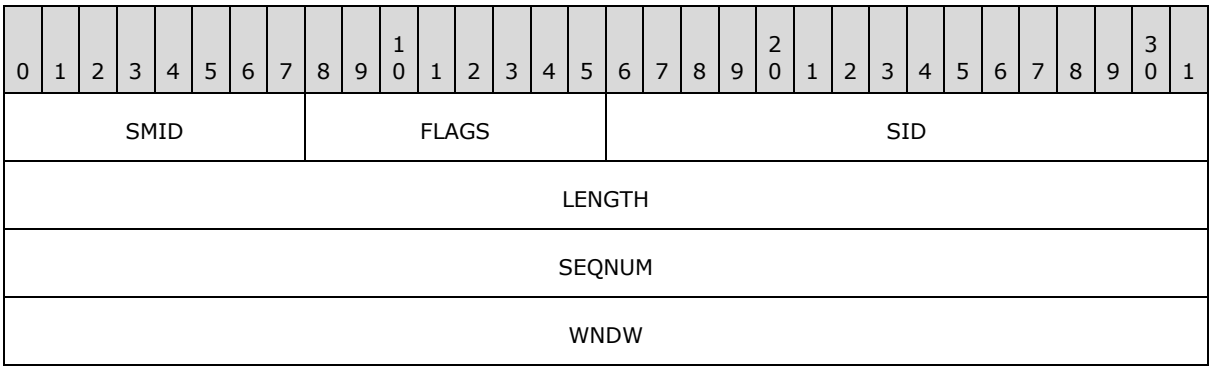
If a SYN packet is received by the client, an error SHOULD be raised to the higher layer and the underlying transport connection SHOULD be closed.

4 Protocol Examples

This section provides examples of SMP packets for various operations being performed.

4.1 Opening a Session

This example illustrates a SYN packet which creates a new session.



SMID (1 byte): 0x53

FLAGS (1 byte): 0x01 (SYN packet)

SID (2 bytes): 0x0000 (The first SMP session on this connection)

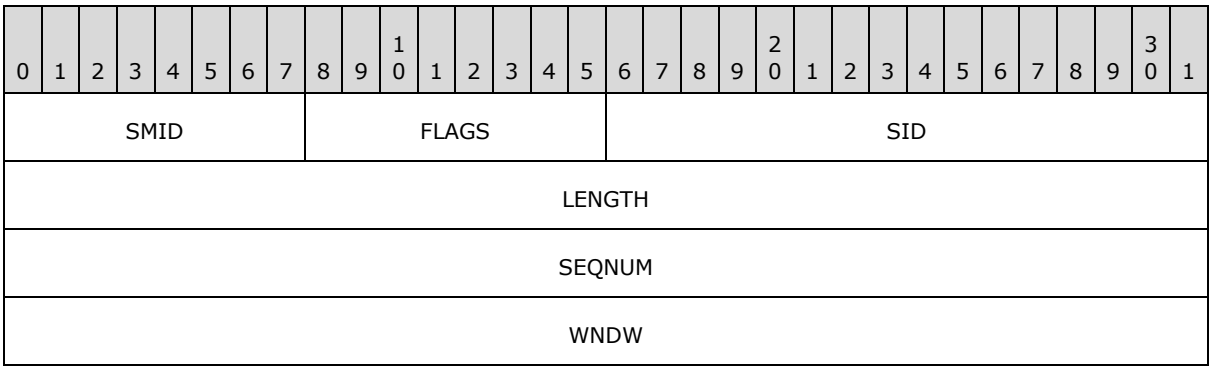
LENGTH (4 bytes): 0x00000010 (The SYN packet does not have any payload)

SEQNUM (4 bytes): 0x00000000 (The initial packet for this session)

WNDW (4 bytes): 0x00000004 (The default of 4 receive buffers posted)

4.2 Update Window - ACK

This example illustrates an ACK packet that updates the peer with a change in window size.



SMID (1 byte): 0x53

FLAGS (1 byte): 0x02 (ACK packet)

SID (2 bytes): 0x0005 (session identifier equals 5)

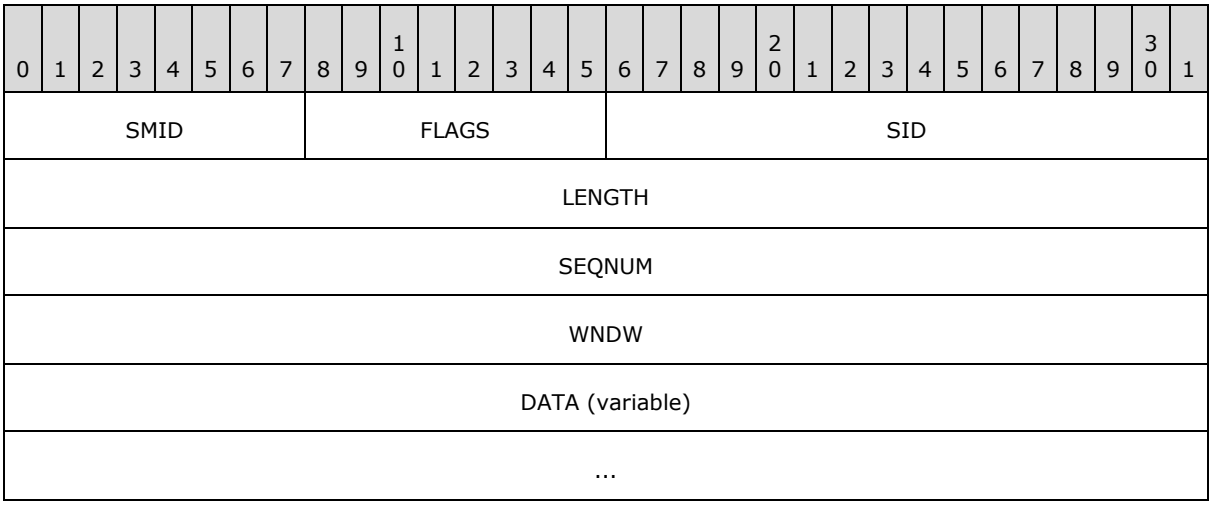
LENGTH (4 bytes): 0x00000010 (The ACK packet does not have a payload)

SEQNUM (4 bytes): 0x00000010

WNDW (4 bytes): 0x00000012

4.3 First Command in a Session

This example illustrates a DATA packet as the first command in a session.



SMID (1 byte): 0x53

FLAGS (1 byte): 0x08 (DATA packet)

SID (2 bytes): 0x0005 (session identifier equals 5)

LENGTH (4 bytes): 0x00000060

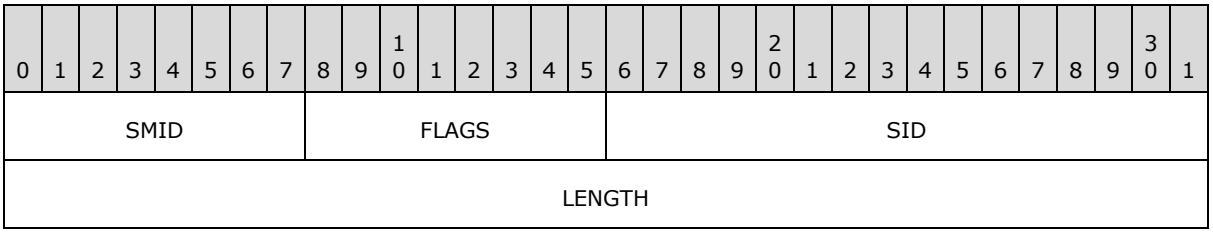
SEQNUM (4 bytes): 0x0000001

WNDW (4 bytes): 0x0000004

DATA (variable): 0x01 01 00 50 00 00 01 00 16 00 00 00 12 00 00 00 02 00 00 00 00 00 00 00 00 00 01 00 00 00 53 00 45 00 54 00 20 00 51 00 55 00 4F 00 54 00 45 00 44 00 5F 00 49 00 44 00 45 00 4E 00 54 00 49 00 46 00 49 00 45 00 52 00 20 00 4F 00 46 00 46 00 (TDS request)

4.4 Closing a Session

This example illustrates the FIN packet as the last command in a session.



SEQNUM
WNDW

SMID (1 byte): 0x53

FLAGS (1 byte): 0x04 (FIN packet)

SID (2 bytes): 0x0005 (session identifier equals 5)

LENGTH (4 bytes): 0x00000010 (The FIN packet does not have a payload)

SEQNUM (4 bytes): 0x0000023

WNDW (4 bytes): 0x0000013

5 Security

5.1 Security Considerations for Implementers

There are no special security considerations for this protocol.

5.2 Index of Security Parameters

None.

6 Appendix A: Product Behavior

The information in this specification is applicable to the following Microsoft products or supplemental software. References to product versions include released service packs:

- Microsoft® SQL Server® 2005
- Microsoft® SQL Server® 2008
- Microsoft® SQL Server® 2008 R2
- Windows Vista® operating system
- Windows Server® 2008 operating system
- Windows® 7 operating system
- Windows Server® 2008 R2 operating system

Exceptions, if any, are noted below. If a service pack or Quick Fix Engineering (QFE) number appears with the product version, behavior changed in that service pack or QFE. The new behavior also applies to subsequent service packs of the product unless otherwise specified. If a product edition appears with the product version, behavior is different in that product edition.

Unless otherwise specified, any statement of optional behavior in this specification that is prescribed using the terms SHOULD or SHOULD NOT implies product behavior in accordance with the SHOULD or SHOULD NOT prescription. Unless otherwise specified, the term MAY implies that the product does not follow the prescription.

<1> Section 3.1.5.2.3:

- When the SMP layer sends a packet, the value of the **LastHighWaterForRecv** ADM variable is set to the value of the **WNDW** field of the sent packet.
- Microsoft products implement a delayed acknowledgement algorithm by sending an ACK packet after every other DATA packet retrieved by the higher layer. In this implementation, an ACK packet is sent if the value of the **HighWaterForRecv** ADM variable minus the value of the **LastHighWaterForRecv** ADM variable is greater than or equal to 2.

7 Change Tracking

No table of changes is available. The document is either new or has had no changes since its last release.

8 Index

R

S

T

U

V